

RAW SEQUENCE LISTING

**The Biotechnology Systems Branch of the Scientific and Technical
Information Center (STIC) no errors detected.**

Application Serial Number: 09/189, 415C
Source: LEW 16
Date Processed by STIC: 08/05/2005

ENTERED



IFW/b

RAW SEQUENCE LISTING

DATE: 08/05/2005

PATENT APPLICATION: US/09/189,415C

TIME: 16:37:08

Input Set : D:\UBCV0004.ST25.txt

Output Set: N:\CRF4\08052005\I189415C.raw

```

3 <110> APPLICANT: Finlay, Brett B.
4      Kenny, Brandant
5      DeVinney, Rebekah
6      Stein, Marcus
8 <120> TITLE OF INVENTION: HOST RECEPTOR FOR PATHOGENIC BACTERIA
10 <130> FILE REFERENCE: UBCV-0004
12 <140> CURRENT APPLICATION NUMBER: US 09/189,415C
13 <141> CURRENT FILING DATE: 1998-11-10
15 <150> PRIOR APPLICATION NUMBER: US 60/065,130
16 <151> PRIOR FILING DATE: 1997-11-12
18 <160> NUMBER OF SEQ ID NOS: 14
20 <170> SOFTWARE: PatentIn version 3.3
22 <210> SEQ ID NO: 1
23 <211> LENGTH: 1920
24 <212> TYPE: DNA
25 <213> ORGANISM: Escherichia coli
27 <400> SEQUENCE: 1
28 cggctgcata ccattacgtc atagtaatat aaaggaacgt gtcaaatttc taaataaaaag      60
30 gatatatgta tgcctattgg taaccttgggt aataatgtaa atggcaatca ttttaattccc      120
32 cctgcgcgcg cactaccttc acaaacagac ggcgcggcac ggggaggaac tgggtcatcta      180
34 attagctcta caggagcatt aggatctcgt tcattgtttt ctcccctgag aaattctatg      240
36 gctgattctg tcgattccag agatattcca ggacttccta caaacccatc gaggcttgct      300
38 gcagctacat ctgagacatg cttgcttgga ggatttgaag ttctccatga taaggggcca      360
40 cttgatattc tcaatacgca aattggaccc tctgcatttc gtgttgaagt gcaggcagat      420
42 ggtactcatg ccgctattgg agaaaaaaat ggtttggagg ttagcgttac attaagtcct      480
44 caagaatgga gcagcttgca atctattgat actgagggtg aaaacagatt tgtttttacc      540
46 gggggacgtg gcggtagtgg gcacccgatg gtcactgtcg catcagatat cgcggaagct      600
48 cgtacgaaaa tactggccaa attagaccca gacaatcatg gaggacgtca acccaaggac      660
50 gttgatacgc gttctgttgg tgttggcagc gtttcgggaa tagatgatgg cgttgtttagc      720
52 gaaacccata cttcaacaac aaattccagc gttcgcctcag atcctaaatt ctggggtttct      780
54 gtcggcgcaa ttgctgctgg tttagcggga ctggcgggaa ctggtattgc acaggcggtg      840
56 gctttgacac cggaaccgga tgatcctaca accaccgatc ctgatcaggc cgcaaattgct      900
58 gcagaaagtg caacaaaaga tcagttaacg caagaagcat tcaagaacct tgagaaccag      960
60 aaagttaaca tcgatgcgaa cggaatgct attccgtctg gggaattaaa agatgatatt      1020
62 gttgagcaaa tagcacaaca agctaaagag gctggtgagg tggccagaca gcaggctgtt      1080
64 gaaagcaatg cacaggcgca gcagcgatat gaggatcagc atgccagacg tcaggaggaa      1140
66 ttacagcttt catcgggtat tggttacggc ctcagcagtg cattgattgt tgctggggga      1200
68 attggtgctg gtgtaacgac tgcgctccat agacgaaatc agccggcaga acagacaact      1260
70 actacaacaa cacatacggg agtgcagcaa cagaccggag ggatacccca gcacaaggtg      1320
72 gcactgatgc cacaagagcg aagacgcttc tctgatagac gtgattcgca ggggagtgtt      1380
74 gcatcgacac actggtcaga ttcctctagc gaagtgggta atccatatgc tgaagtggg      1440
76 ggggctcgga atagtctatc ggctcatcag ccagaagagc atatttatga tgaggctcgt      1500
78 gcagatcctg gttatagcgt tattcagaat ttttcaggga gcggcccagt taccggaagg      1560

```

RAW SEQUENCE LISTING

DATE: 08/05/2005

PATENT APPLICATION: US/09/189,415C

TIME: 16:37:08

Input Set : D:\UBCV0004.ST25.txt

Output Set: N:\CRF4\08052005\I189415C.raw

```

80 ttaataggaa ctccagggca aggtatccaa agtacttatg cgcttctggc aaacagcggc 1620
82 ggattgcgtt taggtatggg aggattaacg agtgggtggc agacggcagt aagttctgta 1680
84 aatgccgcac caacgcaggg accagtacgt ttcgtttaaa tatactctgtg agtatttagt 1740
86 tgaggttggg gtggggtggg ggggcgtttt actagcggtta atgtttcaga gaacaacgtt 1800
88 gcagcatggg taactcttga acttctgtta ttataatcaa ttaagagaaa ttataatgtc 1860
90 atcaagatat gaacttttat tagatagggt tgcggaaaaa attggtgttg gatctatttc 1920
93 <210> SEQ ID NO: 2
94 <211> LENGTH: 549
95 <212> TYPE: PRT
96 <213> ORGANISM: Escherichia coli
99 <220> FEATURE:
100 <221> NAME/KEY: misc_feature
101 <222> LOCATION: (314)..(314)
102 <223> OTHER INFORMATION: Xaa can be any naturally occurring amino acid
104 <400> SEQUENCE: 2
106 Met Pro Ile Gly Asn Leu Gly Asn Asn Val Asn Gly Asn His Leu Ile
107 1 5 10 15
110 Pro Pro Ala Pro Pro Leu Pro Ser Gln Thr Asp Gly Ala Ala Arg Gly
111 20 25 30
114 Gly Thr Gly His Leu Ile Ser Ser Thr Gly Ala Leu Gly Ser Arg Ser
115 35 40 45
118 Leu Phe Ser Pro Leu Arg Asn Ser Met Ala Asp Ser Val Asp Ser Arg
119 50 55 60
122 Asp Ile Pro Gly Leu Pro Thr Asn Pro Ser Arg Leu Ala Ala Ala Thr
123 65 70 75 80
126 Ser Glu Thr Cys Leu Leu Gly Gly Phe Glu Val Leu His Asp Lys Gly
127 85 90 95
130 Pro Leu Asp Ile Leu Asn Thr Gln Ile Gly Pro Ser Ala Phe Arg Val
131 100 105 110
134 Glu Val Gln Ala Asp Gly Thr His Ala Ala Ile Gly Glu Lys Asn Gly
135 115 120 125
138 Leu Glu Val Ser Val Thr Leu Ser Pro Gln Glu Trp Ser Ser Leu Gln
139 130 135 140
142 Ser Ile Asp Thr Glu Gly Lys Asn Arg Phe Val Phe Thr Gly Gly Arg
143 145 150 155 160
146 Gly Gly Ser Gly His Pro Met Val Thr Val Ala Ser Asp Ile Ala Glu
147 165 170 175
150 Ala Arg Thr Arg Ile Leu Ala Lys Leu Asp Pro Asp Asn His Gly Gly
151 180 185 190
154 Arg Gln Pro Lys Asp Val Asp Thr Arg Ser Val Gly Val Gly Ser Ala
155 195 200 205
158 Ser Gly Ile Asp Asp Gly Val Val Ser Glu Thr His Thr Ser Thr Thr
159 210 215 220
162 Asn Ser Ser Val Arg Ser Asp Pro Lys Phe Trp Val Ser Val Gly Ala
163 225 230 235 240
166 Ile Ala Ala Gly Leu Ala Gly Leu Ala Ala Thr Gly Ile Ala Gln Ala
167 245 250 255
170 Leu Ala Leu Thr Pro Glu Pro Asp Asp Pro Thr Thr Thr Asp Pro Asp
171 260 265 270

```

RAW SEQUENCE LISTING

DATE: 08/05/2005

PATENT APPLICATION: US/09/189,415C

TIME: 16:37:08

Input Set : D:\UBCV0004.ST25.txt

Output Set: N:\CRF4\08052005\I189415C.raw

```

174 Gln Ala Ala Asn Ala Ala Glu Ser Ala Thr Lys Asp Gln Leu Thr Gln
175           275           280           285
178 Glu Ala Phe Lys Asn Pro Glu Asn Gln Lys Val Asn Ile Asp Ala Asn
179           290           295           300
W--> 182 Gly Asn Ala Ile Pro Ser Gly Glu Leu Xaa Asp Asp Ile Val Glu Gln
183 305           310           315           320
186 Ile Ala Gln Gln Ala Lys Glu Ala Gly Glu Val Ala Arg Gln Gln Ala
187           325           330           335
190 Val Glu Ser Asn Ala Gln Ala Gln Gln Arg Tyr Glu Asp Gln His Ala
191           340           345           350
194 Arg Arg Gln Glu Glu Leu Gln Leu Ser Ser Gly Ile Gly Tyr Gly Leu
195           355           360           365
198 Ser Ser Ala Leu Ile Val Ala Gly Gly Ile Gly Ala Gly Val Thr Thr
199           370           375           380
202 Ala Leu His Arg Arg Asn Gln Pro Ala Glu Gln Thr Thr Thr Thr Thr
203 385           390           395           400
206 Thr His Thr Val Val Gln Gln Gln Thr Gly Gly Ile Pro Gln His Lys
207           405           410           415
210 Val Ala Leu Met Pro Gln Glu Arg Arg Phe Ser Asp Arg Arg Asp
211           420           425           430
214 Ser Gln Gly Ser Val Ala Ser Thr His Trp Ser Asp Ser Ser Ser Glu
215           435           440           445
218 Val Val Asn Pro Tyr Ala Glu Val Gly Gly Ala Arg Asn Ser Leu Ser
219           450           455           460
222 Ala His Gln Pro Glu Glu His Ile Tyr Asp Glu Val Ala Ala Asp Pro
223 465           470           475           480
226 Gly Tyr Ser Val Ile Gln Asn Phe Ser Gly Ser Gly Pro Val Thr Gly
227           485           490           495
230 Arg Leu Ile Gly Thr Pro Gly Gln Gly Ile Gln Ser Thr Tyr Ala Leu
231           500           505           510
234 Leu Ala Asn Ser Gly Gly Leu Arg Leu Gly Met Gly Gly Leu Thr Ser
235           515           520           525
238 Gly Gly Glu Thr Ala Val Ser Ser Val Asn Ala Ala Pro Thr Pro Gly
239           530           535           540
242 Pro Val Arg Phe Val
243 545
246 <210> SEQ ID NO: 3
247 <211> LENGTH: 1723
248 <212> TYPE: DNA
249 <213> ORGANISM: Escherichia coli
251 <400> SEQUENCE: 3
252 atgcctattg gtaaccttgg tcataatccc aatgtgaata attcaattcc tctgcacct      60
254 ccattacctt cacaaaccga cgggtgcagg gggcggtggtc agctcattaa ctctacgggg      120
256 ccgttgggat ctcggtgcgt atttacgcct gtaaggaatt ctatggctga ttctggcgac      180
258 aatcgtgccg gtgatgttcc tggacttcct gtaaataccga tgcgcctggc ggcgtctgag      240
260 ataacactga atgatggatt tgaagttcct catgatcatg gtccgctcga tactcttaac      300
262 aggcagattg gctcttcggt atttcgagtt gaaactcagg aagatggtaa acatattgct      360
264 gtcggtcaga ggaatggtgt tgagacctct gttgttttaa gtgatcaaga gtacgctcgc      420
266 ttgcagtcca ttgatcctga aggtaaagac aaatttgtat ttactggagg ccgtgggtgt      480

```

RAW SEQUENCE LISTING

DATE: 08/05/2005

PATENT APPLICATION: US/09/189,415C

TIME: 16:37:08

Input Set : D:\UBCV0004.ST25.txt

Output Set: N:\CRF4\08052005\I189415C.raw

```

268 gctgggcatg ctatgggtcac cgttgcttca gatatcacgg aagcccgcca aaggatactg 540
270 gagctgttag agcccaaagg gaccggggag tccaaagggt ctggggagtc aaaaggcggt 600
272 ggggagttga gggagtcaaa tagcgggtgcg gaaaacacca cagaaactca gacctcaacc 660
274 tcaacttcca gccttcgttc agatcctaaa ctttggttgg cgttggggac tgttgctaca 720
276 ggtctgatag ggttggcggc gacgggtatt gtacaggcgc ttgcattgac gccggagccg 780
278 gatagcccaa ccacgaccga ccctgatgca gctgcaagt caactgaaac tgcgacaaga 840
280 gatcagttaa cgaaagaagc gttccagaac ccagataatc aaaaagttaa tatcgatgag 900
282 ctcggaatg cgattccgtc aggggtattg aaagatgatg ttggtgcgaa tatagaagag 960
284 caggctaaag cagcaggcga agaggccaaa cagcaagcca ttgaaaataa tgctcaggcg 1020
286 caaaaaaaat atgatgaaca acaagctaaa cgccaggagg agctgaaagt ttcacgggg 1080
288 gctggctacg gtcttagtgg cgcattgatt cttggtgggg gaattggtgt tgcgcgcacc 1140
290 gctgcgcttc atcgaaaaaa tcagccggtg gaacaaacaa caacaactac tactacaact 1200
292 acaactacaa gcgcacgtac ggtagagaat aagcctgcaa ataatacacc tgcacagggc 1260
294 aatgtagata cccctgggtc agaagatacc atggagagca gacgtagctc gatggctagc 1320
296 acctcgtcga ctttctttga cacttccagc atagggaccg tgcagaatcc gtatgctgat 1380
298 gttaaaacat cgctgcatga ttcgcagggt cgcacttcta attctaatac gtctgttcag 1440
300 aatatgggga atacagattc tgttgatat agcaccattc aacatcctcc ccgggatact 1500
302 actgataacg gcgcacgggt attaggaat ccaagtgcgg ggattcaaag cacttatgcg 1560
304 cgtctggcgc taagtgggtg attacgccat gacatgggag gattaacggg ggggagtaat 1620
306 agcgtctgtg atacttcgaa taaccaccca gcgcgggat cccatcgttt cgtctaaata 1680
308 tatccataat cattttattt agaggggagg aggggggaag tct 1723

```

311 <210> SEQ ID NO: 4

312 <211> LENGTH: 559

313 <212> TYPE: PRT

314 <213> ORGANISM: Escherichia coli

316 <400> SEQUENCE: 4

```

318 Met Pro Ile Gly Asn Leu Gly His Asn Pro Asn Val Asn Asn Ser Ile
319 1 5 10 15
322 Pro Pro Ala Pro Pro Leu Pro Ser Gln Thr Asp Gly Ala Gly Gly Arg
323 20 25 30
326 Gly Gln Leu Ile Asn Ser Thr Gly Pro Leu Gly Ser Arg Ala Leu Phe
327 35 40 45
330 Thr Pro Val Arg Asn Ser Met Ala Asp Ser Gly Asp Asn Arg Ala Ser
331 50 55 60
334 Asp Val Pro Gly Leu Pro Val Asn Pro Met Arg Leu Ala Ala Ser Glu
335 65 70 75 80
338 Ile Thr Leu Asn Asp Gly Phe Glu Val Leu His Asp His Gly Pro Leu
339 85 90 95
342 Asp Thr Leu Asn Arg Gln Ile Gly Ser Ser Val Phe Arg Val Glu Thr
343 100 105 110
346 Gln Glu Asp Gly Lys His Ile Ala Val Gly Gln Arg Asn Gly Val Glu
347 115 120 125
350 Thr Ser Val Val Leu Ser Asp Gln Glu Tyr Ala Arg Leu Gln Ser Ile
351 130 135 140
354 Asp Pro Glu Gly Lys Asp Lys Phe Val Phe Thr Gly Gly Arg Gly Gly
355 145 150 155 160
358 Ala Gly His Ala Met Val Thr Val Ala Ser Asp Ile Thr Glu Ala Arg
359 165 170 175
362 Gln Arg Ile Leu Glu Leu Leu Glu Pro Lys Gly Thr Gly Glu Ser Lys

```

RAW SEQUENCE LISTING

DATE: 08/05/2005

PATENT APPLICATION: US/09/189,415C

TIME: 16:37:08

Input Set : D:\UBCV0004.ST25.txt

Output Set: N:\CRF4\08052005\I189415C.raw

```

363          180          185          190
366 Gly Ala Gly Glu Ser Lys Gly Val Gly Glu Leu Arg Glu Ser Asn Ser
367          195          200          205
370 Gly Ala Glu Asn Thr Thr Glu Thr Gln Thr Ser Thr Ser Thr Ser Ser
371          210          215          220
374 Leu Arg Ser Asp Pro Lys Leu Trp Leu Ala Leu Gly Thr Val Ala Thr
375 225          230          235          240
378 Gly Leu Ile Gly Leu Ala Ala Thr Gly Ile Val Gln Ala Leu Ala Leu
379          245          250          255
382 Thr Pro Glu Pro Asp Ser Pro Thr Thr Thr Asp Pro Asp Ala Ala Ala
383          260          265          270
386 Ser Ala Thr Glu Thr Ala Thr Arg Asp Gln Leu Thr Lys Glu Ala Phe
387          275          280          285
390 Gln Asn Pro Asp Asn Gln Lys Val Asn Ile Asp Glu Leu Gly Asn Ala
391          290          295          300
394 Ile Pro Ser Gly Val Leu Lys Asp Asp Val Val Ala Asn Ile Glu Glu
395 305          310          315          320
398 Gln Ala Lys Ala Ala Gly Glu Glu Ala Lys Gln Gln Ala Ile Glu Asn
399          325          330          335
402 Asn Ala Gln Ala Gln Lys Lys Tyr Asp Glu Gln Gln Ala Lys Arg Gln
403          340          345          350
406 Glu Glu Leu Lys Val Ser Ser Gly Ala Gly Tyr Gly Leu Ser Gly Ala
407          355          360          365
410 Leu Ile Leu Gly Gly Gly Ile Gly Val Ala Val Thr Ala Ala Leu His
411          370          375          380
414 Arg Lys Asn Gln Pro Val Glu Gln Thr Thr Thr Thr Thr Thr Thr Thr
415 385          390          395          400
418 Thr Thr Thr Ser Ala Arg Thr Val Glu Asn Lys Pro Ala Asn Asn Thr
419          405          410          415
422 Pro Ala Gln Gly Asn Val Asp Thr Pro Gly Ser Glu Asp Thr Met Glu
423          420          425          430
426 Ser Arg Arg Ser Ser Met Ala Ser Thr Ser Ser Thr Phe Phe Asp Thr
427          435          440          445
430 Ser Ser Ile Gly Gly Pro Cys Arg Ile Arg Met Leu Met Leu Lys His
431          450          455          460
434 Arg Cys Met Ile Arg Arg Cys Arg Leu Leu Ile Leu Ile Arg Leu Phe
435 465          470          475          480
438 Arg Ile Trp Gly Ile Gln Ile Ser Val Val Tyr Ser Thr Ile Gln His
439          485          490          495
442 Pro Pro Arg Asp Thr Thr Asp Asn Gly Ala Arg Leu Leu Gly Asn Pro
443          500          505          510
446 Ser Ala Gly Ile Gln Ser Thr Tyr Ala Arg Leu Ala Leu Ser Gly Gly
447          515          520          525
450 Leu Arg His Asp Met Gly Gly Leu Thr Gly Gly Ser Asn Ser Ala Val
451          530          535          540
454 Asn Thr Ser Asn Asn Pro Pro Ala Pro Gly Ser His Arg Phe Val
455 545          550          555
458 <210> SEQ ID NO: 5
459 <211> LENGTH: 1460

```

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/09/189,415C

DATE: 08/05/2005
TIME: 16:37:09

Input Set : D:\UBCV0004.ST25.txt
Output Set: N:\CRF4\08052005\I189415C.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:2; Xaa Pos. 314

Invalid <213> Response:

Use of "Artificial" only as "<213> Organism" response is incomplete, per 1.823(b) of New Sequence Rules. Valid response is Artificial Sequence.

Seq#:8,9

VERIFICATION SUMMARY

DATE: 08/05/2005

PATENT APPLICATION: US/09/189,415C

TIME: 16:37:09

Input Set : D:\UBCV0004.ST25.txt

Output Set: N:\CRF4\08052005\I189415C.raw

L:182 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2 after pos.:304